Einar Stole 1814 11/14
305-4507

Serial Number: 08/898,560

CRF Processing Date: 11/12/99
Edited by: _____
Verified by: _____ (STIC staff)

ENTERED

☐ Changed a file from non-ASCII to ASCII

☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.

☐ Edited a format error in the Current Application Data section, specifically: _____

☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____ .

☐ Added the mandatory heading and subheadings for "Current Application Data".

☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.

☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: _____

☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: _____

☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: _____

☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.

☐ Inserted colons after headings/subheadings. Headings edited included: _____

☐ Deleted extra, invalid, headings used by an applicant, specifically: _____

☐ Deleted: ☐ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as _____ .

☐ Inserted mandatory headings, specifically: _____

☐ Corrected an obvious error in the response, specifically: _____

☐ Edited identifiers where upper case is used but lower case is required, or vice versa.

☐ Corrected an error in the Number of Sequences field, specifically: _____

☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.

☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: _____

☑ Other: Seqr 1,2 - changed "KEY" to "NAME/KEY" under (ix)FEATURE
_____
_____

Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

3/1/95

1814

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

DATE: 11/12/97
TIME: 21:24:16

INPUT SET: S21525.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

```
 1                          SEQUENCE LISTING
 2
 3   (1)    General Information:
 4
 5   (i) APPLICANT: Hiroyuki NAKANE,  Chikara OHTO,  Shinichi OHNUMA,
 6                  Kazutake HIROOKA,  Tokuzo NISHINO
 7
 8   (ii) TITLE OF INVENTION: Farnesyl Diphosphate Synthase
 9
10   (iii) NUMBER OF SEQUENCES: 14
11
12   (iv) CORRESPONDENCE ADDRESS:
13      (A) ADDRESSEE: Kenyon & Kenyon
14      (B) STREET: One Broadway
15      (C) CITY: New York
16      (D) STATE: NY
17      (E) COUNTRY: USA
18      (F) ZIP: 10004
19
20   (v) COMPUTER READABLE FORM:
21      (A) MEDIUM TYPE: 3" Floppy disk
22      (B) COMPUTER: IBM PC compatible
23      (C) OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
24      (D) SOFTWARE: WordPerfect 6.1 Windows
25
26   (vi) CURRENT APPLICATION DATA:
27      (A) APPLICATION NUMBER: Not Yet Issued
28      (B) FILING DATE: Concurrent Herewith
29      (C) CLASSIFICATION: Not Yet Issued
30
31   (vii) PRIOR APPLICATION DATA:
32      (A) APPLICATION NUMBER: JP 8-213211
33      (B) FILING DATE: 24-JUL-96
34
35   (viii) ATTORNEY/AGENT INFORMATION:
36      (A) NAME: Edward W. Greason
37      (B) REGISTRATION NUMBER: 18,918
38      (C) REFERENCE/DOCKET NUMBER: 77670/495
39
40   (ix) TELECOMMUNICATION INFORMATION:
41      (A) TELEPHONE: (212)425-7200
42      (B) TELEFAX: (212)425-5288
43
44      (2) INFORMATION FOR SEQ ID NO:1:
45
46   (i) SEQUENCE CHARACTERISTICS:
```

--> 27
--> 29

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

DATE: 11/12/97
TIME: 21:24:19

INPUT SET: S21525.raw

```
47        (A) LENGTH: 330 amino acids
48        (B) TYPE: amino acid
49        (D) TOPOLOGY: linear
50
51    (ii) MOLECULE TYPE: protein
52
53    (vi) ORIGINAL SOURCE:
54        (A) ORGANISM: Sulfolobus acidocaldarius
55        (B) STRAIN: ATCC 33909
56
57    (ix) FEATURE:
58        (A) NAME/KEY: Asp-rich domain
59        (B) LOCATION: 82-86
60
61    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
62    .
63    Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn
64                    5                   10                  15
65    Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu
66                   20                   25                  30
67    Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu
68                   35                   40                  45
69    Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala
70         50                   55                  60
71    Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val
72    65                   70                  75                  80
73    His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr
74                   85                   90                  95
75    Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu
76                  100                  105                 110
77    Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu
78                  115                  120                 125
79    Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile
80        130                  135                 140
81    Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg
82    145                  150                 155                 160
83    Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr
84                  165                  170                 175
85    Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly
86                  180                  185                 190
87    Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu
88                  195                  200                 205
89    Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp
90        210                  215                 220
91    Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
92    225                  230                 235                 240
93
94    Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu
95                  245                  250                 255
96    Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu
97                  260                  265                 270
98    Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
99        275                  280                 285
```

INPUT SET: S21525.raw

```
100    Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
101        290                 295                 300
102    Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
103    305                 310                 315                 320
104    Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
105                    325                 330
106
107        (2) INFORMATION FOR SEQ ID NO:2:
108
109    (i) SEQUENCE CHARACTERISTICS:
110        (A) LENGTH: 993 base pairs
111        (B) TYPE: nucleic acid
112        (C) STRANDEDNESS: double
113        (D) TOPOLOGY: linear
114
115    (ii) MOLECULE TYPE: genomic DNA
116
117    (vi) ORIGINAL SOURCE:
118        (A) ORGANISM: Sulfolobus acidocaldarius
119        (B) STRAIN: ATCC 33909
120
121    (ix) FEATURE:
122        (A) NAME/KEY: Asp-rich domain coding
123        (B) LOCATION: 246-258
124
125    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
126
127    ATGAGTTACT TTGACAACTA TTTTAATGAG ATTGTTAATT CTGTAAACGA CATTATTAAG    60
128
129    AGCTATATAT CTGGAGATGT TCCTAAACTA TATGAAGCCT CATATCATTT GTTTACATCT    120
130
131    GGAGGTAAGA GGTTAAGACC ATTAATCTTA ACTATATCAT CAGATTTATT CGGAGGACAG    180
132
133    AGAGAAAGAG CTTATTATGC AGGTGCAGCT ATTGAAGTTC TTCATACTTT TACGCTTGTG    240
134
135    CATGATGATA TTATGGATCA AGATAATATC AGAAGAGGGT TACCCACAGT CCACGTGAAA    300
136
137    TACGGCTTAC CCTTAGCAAT ATTAGCTGGG GATTTACTAC ATGCAAAGGC TTTTCAGCTC    360
138
139    TTAACCCAGG CTCTTAGAGG TTTGCCAAGT GAAACCATAA TTAAGGCTTT CGATATTTTC    420
140
141    ACTCGTTCAA TAATAATTAT ATCCGAAGGA CAGGCAGTAG ATATGGAATT TGAGGACAGA    480
142
143    ATTGATATAA AGGAGCAGGA ATACCTTGAC ATGATCTCAC GTAAGACAGC TGCATTATTC    540
144
145    TCGGCATCCT CAAGTATAGG CGCACTTATT GCTGGTGCTA ATGATAATGA TGTAAGACTG    600
146
147    ATGTCTGATT TCGGTACGAA TCTAGGTATT GCATTTCAGA TTGTTGACGA TATCTTAGGT    660
148
149    CTAACAGCAG ACGAAAAGGA ACTTGGAAAG CCTGTTTTTA GTGATATTAG GGAGGGTAAA    720
150
151    AAGACTATAC TTGTAATAAA AACACTGGAG CTTTGTAAAG AGGACGAGAA GAAGATTGTC    780
152
```

INPUT SET: S21525.raw

153   CTAAAGGCGT TAGGTAATAA GTCAGCCTCA AAAGAAGAAT TAATGAGCTC AGCAGATATA   840
154
155   ATTAAGAAAT ACTCTTTAGA TTATGCATAC AATTTAGCAG AGAAATATTA TAAAAATGCT   900
156
157   ATAGACTCTT TAAATCAAGT CTCCTCTAAG AGTGATATAC CTGGAAAGGC TTTAAAATAT   960
158
159   CTAGCTGAAT TTACGATAAG AAGGAGAAAA TAA   993
160
161
162   (2) INFORMATION FOR SEQ ID NO:3:
163
164   (i) SEQUENCE CHARACTERISTICS:
165   (A) LENGTH:37
166   (B) TYPE: nucleic acid
167   (C) STRANDEDNESS: single
168   (D) TOPOLOGY: linear
169
170   (ii) MOLECULE TYPE: cDNA
171
172   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
173
174   CATACTTTTT TCCTTGTGGC TGATGATATC ATGGATC   37
175
176
177   (2) INFORMATION FOR SEQ ID NO:4:
178
179   (i) SEQUENCE CHARACTERISTICS:
180   (A) LENGTH:37
181   (B) TYPE: nucleic acid
182   (C) STRANDEDNESS: single
183   (D) TOPOLOGY: linear
184
185   (ii) MOLECULE TYPE: cDNA
186
187   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
188
189   CATACTTTTT TCCTTGTGCT TGATGATATC ATGGATC   37
190
191   (2) INFORMATION FOR SEQ ID NO:5:
192
193   (i) SEQUENCE CHARACTERISTICS:
194   (A) LENGTH:37
195   (B) TYPE: nucleic acid
196   (C) STRANDEDNESS: single
197   (D) TOPOLOGY: linear
198
199   (ii) MOLECULE TYPE: cDNA
200
201   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
202
203   CATACTTATT TCCTTGTGCT TGATGATATC ATGGATC   37
204
205

```
206        (2) INFORMATION FOR SEQ ID NO:6:
207
208    (i) SEQUENCE CHARACTERISTICS:
209        (A) LENGTH:37
210        (B) TYPE: nucleic acid
211        (C) STRANDEDNESS: single
212        (D) TOPOLOGY: linear
213
214    (ii) MOLECULE TYPE: cDNA
215
216    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
217
218    CATACTTATT TCCTTGTGGC TGATGATATC ATGGATC              37
219
220
221        (2) INFORMATION FOR SEQ ID NO:7:
222
223    (i) SEQUENCE CHARACTERISTICS:
224        (A) LENGTH: 36
225        (B) TYPE: nucleic acid
226        (C) STRANDEDNESS: single
227        (D) TOPOLOGY: linear
228
229    (ii) MOLECULE TYPE: cDNA
230
231
232    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
233
234    GTTCTTCATA CTTATTCGCT TATTCATGAT AGTATT               36
235
236
237
238
239        (2) INFORMATION FOR SEQ ID NO:8:
240
241    (i) SEQUENCE CHARACTERISTICS:
242        (A) LENGTH: 33
243        (B) TYPE: nucleic acid
244        (C) STRANDEDNESS: single
245        (D) TOPOLOGY: linear
246
247    (ii) MOLECULE TYPE: cDNA
248
249    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
250
251    ATTCATGATG ATCTTCCATC GATGGATCAA GAT                  33
252
253
254        (2) INFORMATION FOR SEQ ID NO:9:
255
256    (i) SEQUENCE CHARACTERISTICS:
257        (A) LENGTH: 27
258        (B) TYPE: nucleic acid
```

PAGE: 6

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

DATE: 11/12/97
TIME: 21:24:33

INPUT SET: S21525.raw

***** PREVIOUSLY ERRORED SEQUENCES - EDITED *****

```
44        (2) INFORMATION FOR SEQ ID NO:1:
45
46    (i) SEQUENCE CHARACTERISTICS:
47        (A) LENGTH: 330 amino acids
48        (B) TYPE: amino acid
49        (D) TOPOLOGY: linear
50
51    (ii) MOLECULE TYPE: protein
52
53    (vi) ORIGINAL SOURCE:
54        (A) ORGANISM: Sulfolobus acidocaldarius
55        (B) STRAIN: ATCC 33909
56
57    (ix) FEATURE:
58        (A) NAME/KEY: Asp-rich domain
59        (B) LOCATION: 82-86
60
61    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
62
63    Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn
64                    5                  10                  15
65    Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu
66                   20                  25                  30
67    Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu
68                   35                  40                  45
69    Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala
70            50                  55                  60
71    Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val
72    65                  70                  75                  80
73    His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr
74                   85                  90                  95
75    Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu
76                  100                 105                 110
77    Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu
78                  115                 120                 125
79    Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile
80            130                 135                 140
81    Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg
82    145                 150                 155                 160
83    Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr
84                  165                 170                 175
85    Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly
86                  180                 185                 190
87    Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu
88                  195                 200                 205
89    Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp
90            210                 215                 220
```

INPUT SET: S21525.raw

```
 91    Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
 92    225             230             235             240
 93
 94    Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu
 95                    245             250             255
 96    Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu
 97                260             265             270
 98    Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
 99        275             280             285
100    Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
101        290             295             300
102    Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
103    305             310             315             320
104    Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
105                325             330
106
```

107   (2) INFORMATION FOR SEQ ID NO:2:
108
109   (i) SEQUENCE CHARACTERISTICS:
110   (A) LENGTH: 993 base pairs
111   (B) TYPE: nucleic acid
112   (C) STRANDEDNESS: double
113   (D) TOPOLOGY: linear
114
115   (ii) MOLECULE TYPE: genomic DNA
116
117   (vi) ORIGINAL SOURCE:
118   (A) ORGANISM: Sulfolobus acidocaldarius
119   (B) STRAIN: ATCC 33909
120
121   (ix) FEATURE:
122   (A) NAME/KEY: Asp-rich domain coding
123   (B) LOCATION: 246-258
124
125   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
126
127   ATGAGTTACT TTGACAACTA TTTTAATGAG ATTGTTAATT CTGTAAACGA CATTATTAAG   60
128
129   AGCTATATAT CTGGAGATGT TCCTAAACTA TATGAAGCCT CATATCATTT GTTTACATCT   120
130
131   GGAGGTAAGA GGTTAAGACC ATTAATCTTA ACTATATCAT CAGATTTATT CGGAGGACAG   180
132
133   AGAGAAAGAG CTTATTATGC AGGTGCAGCT ATTGAAGTTC TTCATACTTT TACGCTTGTG   240
134
135   CATGATGATA TTATGGATCA AGATAATATC AGAAGAGGGT TACCCACAGT CCACGTGAAA   300
136
137   TACGGCTTAC CCTTAGCAAT ATTAGCTGGG GATTTACTAC ATGCAAAGGC TTTTCAGCTC   360
138
139   TTAACCCAGG CTCTTAGAGG TTTGCCAAGT GAAACCATAA TTAAGGCTTT CGATATTTTC   420
140
141   ACTCGTTCAA TAATAATTAT ATCCGAAGGA CAGGCAGTAG ATATGGAATT TGAGGACAGA   480
142

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

INPUT SET: S21525.raw

```
143   ATTGATATAA AGGAGCAGGA ATACCTTGAC ATGATCTCAC GTAAGACAGC TGCATTATTC   540
144
145   TCGGCATCCT CAAGTATAGG CGCACTTATT GCTGGTGCTA ATGATAATGA TGTAAGACTG   600
146
147   ATGTCTGATT TCGGTACGAA TCTAGGTATT GCATTTCAGA TTGTTGACGA TATCTTAGGT   660
148
149   CTAACAGCAG ACGAAAAGGA ACTTGGAAAG CCTGTTTTTA GTGATATTAG GGAGGGTAAA   720
150
151   AAGACTATAC TTGTAATAAA AACACTGGAG CTTTGTAAAG AGGACGAGAA GAAGATTGTC   780
152
153   CTAAAGGCGT TAGGTAATAA GTCAGCCTCA AAAGAAGAAT TAATGAGCTC AGCAGATATA   840
154
155   ATTAAGAAAT ACTCTTTAGA TTATGCATAC AATTTAGCAG AGAAATATTA TAAAAATGCT   900
156
157   ATAGACTCTT TAAATCAAGT CTCCTCTAAG AGTGATATAC CTGGAAAGGC TTTAAAATAT   960
158
159   CTAGCTGAAT TTACGATAAG AAGGAGAAAA TAA                                 993
160
161
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/898,560*

INPUT SET: S21525.raw

Line	Error	Original Text
27	Wrong application Serial Number	(A) APPLICATION NUMBER: Not Yet Issued
29	Wrong Classification	(C) CLASSIFICATION: Not Yet Issued

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

DATE: 11/12/97
TIME: 16:24:22

INPUT SET: S21525.raw

> This Raw Listing contains the General
> Information Section and those Sequences
> containing ERRORS.

SEQUENCE LISTING   *Does Not Comply*
Corrected Diskette Needed

```
 1
 2
 3    (1)     General Information:
 4
 5    (i) APPLICANTS: Hiroyuki NAKANE,   Chikara OHTO,   Shinichi OHNUMA,
 6    Kazutake HIROOKA,   Tokuzo NISHINO
 7
 8    (ii) TITLE OF INVENTION: Farnesyl Diphosphate Synthase
 9
10    (iii) NUMBER OF SEQUENCES: 14
11
12    (iv) CORRESPONDENCE ADDRESS:
13       (A) ADDRESSEE: Kenyon & Kenyon
14       (B) STREET: One Broadway
15       (C) CITY: New York
16       (D) STATE: NY
17       (E) COUNTRY: USA
18       (F) ZIP: 10004
19
20    (v) COMPUTER READABLE FORM:
21       (A) MEDIUM TYPE: 3" Floppy disk
22       (B) COMPUTER: IBM PC compatible
23       (C) OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
24       (D) SOFTWARE: WordPerfect 6.1 Windows
25
26    (vi) CURRENT APPLICATION DATA:
27       (A) APPLICATION NUMBER: Not Yet Issued
28       (B) FILING DATE: Concurrent Herewith
29       (C) CLASSIFICATION: Not Yet Issued
30
31    (vii) PRIOR APPLICATION DATA:
32       (A) APPLICATION NUMBER: JP 8-213211
33       (B) FILING DATE: 24-JUL-96
34
35    (viii) ATTORNEY/AGENT INFORMATION:
36       (A) NAME: Edward W. Greason
37       (B) REGISTRATION NUMBER: 18,918
38       (C) REFERENCE/DOCKET NUMBER: 77670/495
39
40    (ix) TELECOMMUNICATION INFORMATION:
41       (A) TELEPHONE: (212)425-7200
42       (B) TELEFAX: (212)425-5288
43
```

--> α((lines 27)
--> oll (line 29)

ERRORED SEQUENCES FOLLOW:

Best Available Copy

PAGE 2

RAW SEQUENCE LISTING
PATENT APPLICATION   US/08/898,560

DATE: 11/12/97
TIME: 16:24:25

INPUT SET: S21525.raw

44   (2) INFORMATION FOR SEQ ID NO:1:

45

46   (i) SEQUENCE CHARACTERISTICS:

47   (A) LENGTH: 330 amino acids

48   (B) TYPE: amino acid

49   (D) TOPOLOGY: linear

50

51   (ii) MOLECULE TYPE: protein

52

53   (vi) ORIGINAL SOURCE:

54   (A) ORGANISM: Sulfolobus acidocaldarius

55   (B) STRAIN: ATCC 33909

56

57   (ix) FEATURE: _____) *NAME/KEY:*

58   (A) KEY Asp-rich domain

59   (B) LOCATION: 82-86

60

61   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

62

63   Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn

64   5   10   15

65   Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu

66   20   25   30

67   Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu

68   35   40   45

69   Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala

70   50   55   60

71   Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val

72   65   70   75   80

73   His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr

74   85   90   95

75   Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu

76   100   105   110

77   Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu

78   115   120   125

79   Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile

80   130   135   140

81   Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg

82   145   150   155   160

83   Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr

84   165   170   175

85   Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly

86   180   185   190

87   Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu

88   195   200   205

89   Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp

90   210   215   220

91   Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys

92   225   230   235   240

93

94   Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu

95   245   250   255

96   Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu

```
 97                   260                      265                           270
 98     Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
 99                   275                      280                           285
100     Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
101            290                      295                           300
102     Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
103     305                      310                      315                320
104     Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
105                      325                      330
106
107         (2) INFORMATION FOR SEQ ID NO:2:
108
109     (i) SEQUENCE CHARACTERISTICS:
110         (A) LENGTH: 993 base pairs
111         (B) TYPE: nucleic acid
112         (C) STRANDEDNESS: double
113         (D) TOPOLOGY: linear
114
115     (ii) MOLECULE TYPE: genomic DNA
116
117     (vi) ORIGINAL SOURCE:
118         (A) ORGANISM: Sulfolobus acidocaldarius
119         (B) STRAIN: ATCC 33909
120
121     (ix) FEATURE:
122         (A) KEY: Asp-rich domain coding
123         (B) LOCATION: 246-258
124
125     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
126
127     ATGAGTTACT TTGACAACTA TTTTAATGAG ATTGTTAATT CTGTAAACGA CATTATTAAG    60
128
129     AGCTATATAT CTGGAGATGT TCCTAAACTA TATGAAGCCT CATATCATTT GTTTACATCT   120
130
131     GGAGGTAAGA GGTTAAGACC ATTAATCTTA ACTATATCAT CAGATTTATT CGGAGGACAG   180
132
133     AGAGAAAGAG CTTATTATGC AGGTGCAGCT ATTGAAGTTC TTCATACTTT TACGCTTGTG   240
134
135     CATGATGATA TTATGGATCA AGATAATATC AGAAGAGGGT TACCCACAGT CCACGTGAAA   300
136
137     TACGGCTTAC CCTTAGCAAT ATTAGCTGGG GATTTACTAC ATGCAAAGGC TTTTCAGCTC   360
138
139     TTAACCCAGG CTCTTAGAGG TTTGCCAAGT GAAACCATAA TTAAGGCTTT CGATATTTTC   420
140
141     ACTCGTTCAA TAATAATTAT ATCCGAAGGA CAGGCAGTAG ATATGGAATT TGAGGACAGA   480
142
143     ATTGATATAA AGGAGCAGGA ATACCTTGAC ATGATCTCAC GTAAGACAGC TGCATTATTC   540
144
145     TCGGCATCCT CAAGTATAGG CGCACTTATT GCTGGTGCTA ATGATAATGA TGTAAGACTG   600
146
147     ATGTCTGATT TCGGTACGAA TCTAGGTATT GCATTTCAGA TTGTTGACGA TATCTTAGGT   660
148
```

RAW SEQUENCE LISTING
PATENT APPLICATION   US/08/898,560

INPUT SET: S21525.raw

```
149   CTAACAGCAG ACGAAAAGGA ACTTGGAAAG CCTGTTTTTA GTGATATTAG GGAGGGTAAA   720
150
151   AAGACTATAC TTGTAATAAA AACACTGGAG CTTTGTAAAG AGGACGAGAA GAAGATTGTC   780
152
153   CTAAAGGCGT TAGGTAATAA GTCAGCCTCA AAAGAAGAAT TAATGAGCTC AGCAGATATA   840
154
155   ATTAAGAAAT ACTCTTTAGA TTATGCATAC AATTTAGCAG AGAAATATTA TAAAAATGCT   900
156
157   ATAGACTCTT TAAATCAAGT CTCCTCTAAG AGTGATATAC CTGGAAAGGC TTTAAAATAT   960
158
159   CTAGCTGAAT TTACGATAAG AAGGAGAAAA TAA                                993
160
161
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/898,560*

DATE: 11/12/97
TIME: 16:24:35

INPUT SET: S21525.raw

Line	Error	Original Text
27	Wrong application Serial Number	(A) APPLICATION NUMBER: Not Yet Issued
29	Wrong Classification	(C) CLASSIFICATION: Not Yet Issued
58	Unknown or Misplaced Identifier	(A) KEY: Asp-rich domain
122	Unknown or Misplaced Identifier	(A) KEY: Asp-rich domain coding